Internet Backgrounder

The Internet and the World-Wide Web are hot, hot, hot. So hot that we are all a little sick-n-tired of hearing about them. The fact is, feature articles on the World-Wide Web, ground-breaking Web browsers, Netscape's phenomal growth, Microsoft's answer to Netscape's growth and other articles have appeared in The Wall Street Journal and The New York Times. Both of these newspapers have a distinquished presence on the World-Wide Web. Journals as diverse as Rolling Stone and Business Week are singing the praises of the Internet's geometric growth, the use of the intranet as an answer to corporate headaches and anarchic nature of the technology push. The shelves at Barnes & Noble, B. Dalton, Brentanos, Walden, and other chain bookstores are bent to breaking under heaps of opportunistic titles like; *"The Internet for Dummies"*, *"The Idiot's Guide to the Internet"*, *"Free Stuff from the Internet"*, *"E-Mail Addresses of the Rich and Famous"*, and "*Meeting People on the Internet*". Even Public Broadcasting System in the United States has sanctified the Internet with a portentous television special or two.

This technology push of the Internet grew from a Department of Defense experiment in the early 1960s to a world-wide mesh of millions of computers and tens of thousands of networks. Of course, the mechanical aspects of information publishing on the Internet have evolved greatly over that period, from the terminal-like interfaces of the early Telnet servers and network news, to FTP file archive servers, to text menu–based Gopher servers, to the multimedia, bandwidth-devouring World-Wide Web servers of today. But the basic goals have not changed at all. The World Wide Web was designed back in 1989 by Tim Berners-Lee, a computer scientist working at the CERN particle physics lab in Switzerland, as a mechanism to help an international group of physicists collaborate on-line in their research. It came into its own about 1993 or so, when the Web's originally UNIX-based NCSA Mosaic browser hit the desktops of mainstream Windows and Mac users and sent them adrift on the Internet. Up to this point, internet users were mainly scientific and academic 'geeks' using such interfaces as FTP and Telnet protocols.

The intriguing deal about the Internet is that it is a level playing field. If you can get yourself on the Internet, you too can set up your own World-Wide Web server and become an instant electronic information publishing presence. Your work will stand or fall entirely on its own merits--it will be just as visible and accessible to the worldwide community as the material published on the Internet by the White House, Ziff-Davis, Microsoft, the Louvre, or any of the other bastions of power, wealth, and culture found in our society. Within the broadest limits imaginable, there is no one to really stop you, no one to second-guess you, no one to totally censor you; in short, no one to interfere with your ability to reach out to potentially millions of like-minded "citizens of the net" anywhere in the world.

What you need to know about TCP/IP and the Internet

Most of us know of *TCP/IP* as the glue that binds the Internet. The reason for this paper grew out of a number of good questions and good ideas from you asking for a better description of what TCP/IP is and how it works is the reason for this paper. Understanding TCP/IP is largely a matter of understanding a collection of arcane protocols that TCP/IP hosts use to exchange information. TCP/IP stands for *Transmission Control Protocol/Internet Protocol*. In networking terms, it is a protocol. A protocol is an agreed-upon standard connection method permitting two computers to exchange data. TCP/IP isn't just one protocol, but several. That's why you'll often hear it referred to as a protocol suite, of which *TCP* and *IP* are the two principal protocols. The TCP/IP software installed on an Intel-based server, workstation, or desktop computer provides platform-specific implementations of TCP, IP, and other members of the *TCP/IP family*.

The Internet's underlying data handler is the Transmission Control Protocol/Internet Protocol. TCP/IP earned it s place on the Internet with its ability to connect dissimilar computer systems. TCP/IP breaks the data into packets, then routes those packets to their destination along whatever channel is available. There is no guarantee that two successive packets will follow the same pathway, or that they will arrive at their destination in the order transmitted. When the network becomes congested, TCP/IP implements a basic form of flow control. It simply throws away packets when there is no room for them and requests

the originator to resend the information. The protocol calls for the transmitting node to recognize an excess of retry requests and slow its transmission rate to better match the available *bandwidth*. The scheme works well when moving blocks of static information. Out-of-sequence packets pose no problem; file transfer operations simply store incoming data and reassemble it once it has all arrived. Similarly, the dropping and resending of packets is not a problem. Because the operation forwards data to the application only after all the data is in, the effect of packet dropping is merely to delay the transfer's completion.

Network designers often use the seven-layer ISO/OSI (International Standards Organization/Open Systems Interconnect) model when discussing *TCP/IP architectures*. Each layer in the model corresponds to one level of network functionality. At the bottom sits the physical layer, which represents the physical medium through which data travels--in other words, the network cabling. Above that is the data-link layer, whose services are provided by network interface cards. The uppermost layer is the application layer, where application programs that use network services run.

Current TCP/IP Status

At the desktop, TCP/IP finally surpassed Novell NetWare's IPX as the LAN transport protocol of choice. According to a research report by the Business Research Group (BRG), 1995 was the year that TCP/IP seized the desktop at large and medium-sized corporations in the United States. According to the report, TCP/IP will account for 46% of LAN traffic, while IPX will account for 43%. In the server market segment, TCP/IP commands a huge following as the LAN protocol of choice according to a report by CAP Ventures.

The historical affinity of *Ethernet* and TCP/IP--and Ethernet's enduring *scalability*--seems certain to cement Ethernet's place as the *LAN* access hardware of choice for the near term. The Business Research Group reports in 1995, some 40% of corporations were using switched 100Mbit/second Ethernet. By comparison, approximately 31% of firms use the less expensive shared 100Mbit/second Ethernet. Many industry leaders believe switched 100Mbit/second Ethernet will capture an even higher proportion of the 1997 network market share segment as its price gets even more palatable. Token Ring sales, by contrast, are being totally displaced. A research report by the San Jose market research firm Dataquest last year found that by 1999, only 5.8% of all LAN types shipped will be Token Ring systems--down from around 25% in 1994.

For IS managers, the solidifying of a transport protocol standard, the affirmation of Ethernet and the apparent stalling of ATM have simplified choices. TCP/IP and Ethernet are familiar technologies. In the data communication hardware market, the division between LANs and WANs is crumbling like the Berlin Wall. The increasing homogeneity of communications technologies and the consolidation of key industry players are combining to remake the networked landscape. Moreover, IS managers have been able to forestall the high cost of upgrading corporate networks to full-blown ATM--not long ago hyped as the next backbone networking technology of choice.

The impending conversion of the cable TV companies and the deregulated telephone companies into network service providers will accelerate this growth of the Internet even further. Within a few years, you can confidently expect that every household and every small business will have some form of high-bandwidth digital connectivity to the Internet.

Summary

- TCP/IP is a set of protocols that permit physical networks to be joined together to form an internet. TCP/IP combines the individual networks to form a virtual network in which individual hosts are identified not by physical network addresses but by IP addresses.
- TCP/IP is the protocol which is the backbone and driving force that propels the global Internet. In fact, it is to the Internet that TCP/IP owes much of its popularity.

- TCP/IP uses a multilayered architecture that clearly defines each protocol's responsibilities. TCP and *UDP* provide high-level data transmission services to network application programs, and both rely on IP to transmit packets of data. IP is responsible for routing the packets to their destination.
- Data moving between two applications running on Internet hosts travels up and down the hosts' TCP/IP stacks.
- Information added by the TCP/IP modules on the sending end is stripped off by the corresponding TCP/IP modules on the receiving end and used to re-create the original data.
- The R440LX baseboard contains an embedded *SCSI-2 (wide)* controller, an embedded video controller and best of all… the industry leading 10/100Mbit network controller… the Intel 82557 NIC.

TCP/IP Information Bubble: TCP/IP is used with networked systems to communicate with each other with the TCP/IP protocol, it really doesn't matter whether the systems are part of the same network or are attached to separate networks. One machine can be a mainframe computer (i.e. IBM AS400) and the others can be Intel-based personal computers. TCP/IP is a platform-independent standard that bridges the gap between dissimilar computers, operating systems, and networks. TCP/IP is an outgrowth of research funded by the U.S. government's Advanced Research Projects Agency (ARPA) in the 1970s. Development was started so research networks around the world could be joined to form a virtual network known as an *internetwork*. The original Internet was formed by converting an existing conglomeration of networks, known as *ARPAnet*, over to TCP/IP…that ARPAnet concept with the TCP/IP protocol would eventually become the backbone of today's Internet.

The reason TCP/IP is so important today is that it allows standalone networks to be connected to the Internet or linked together to create private intranets. TCP/IP provides a solution to the problem of how two computers attached to the same intranet, but belonging to different physical networks can exchange data. The solution comes in several parts, with each member of the TCP/IP protocol suite filling in one piece of the puzzle. The networks that comprise an intranet are physically connected by devices called *routers* or IP routers. On a TCP/IP intranet, information travels in discrete units called IP packets or IP datagrams. The most fundamental TCP/IP protocol, *IP*, transmits Internet Protocol *datagrams* across an intranet and performs an important function called routing --choosing the path that datagrams will follow to get from point A to point B and using routers to basically *hop* from network to network. TCP is a higher-level protocol that allows applications running on different hosts to exchange data streams. TCP divides data streams into chunks called TCP segments and transmits them using IP. In most cases, each *TCP segment* is sent in a single IP datagram. If necessary, however, TCP will split segments into multiple IP datagrams that are compatible with the physical data frames that carry bits and bytes between hosts on a network. Because IP doesn't guarantee that datagrams will be received in the same order in which they were sent, TCP reassembles TCP segments at the other end to form an uninterrupted data stream. *FTP* and telnet are two examples of popular TCP/IP applications that rely on TCP. TCP/IP software makes each PC system attached to the network a sibling to all the others. The protocol software hides the routers and underlying network architectures and makes the whole conglomerate seem like one big network. Just as connections to an Ethernet network are identified by 48-bit *Ethernet* IDs, connections to an intranet are identified by 32-bit *IP addresses*, which we express as dotted decimal numbers (for example, 128.10.2.3). Given a remote computer's IP address, a computer on an intranet (or the Internet) can send data to the remote computer as if the two were part of the same physical network.

TCP/IP Family member info bubble: Another important member of the TCP/IP suite is the *User Datagram Protocol (UDP*), which is similar to but more primitive than TCP. TCP is a ‘reliable’ protocol because it performs the error-checking and handshaking necessary to verify that data makes it to its destination intact. UDP is an ‘unreliable’ protocol because it doesn't guarantee that datagrams will arrive in the order in which they were sent or even that they will arrive at all. If reliability is desired, it's up to the application to provide it. Still, UDP has its place in the TCP/IP universe, and a number of

applications use it. The *SNMP (Simple Network Management Protocol)* application provided with most implementations of TCP/IP is one example of a UDP application.

TCP/IP Architecture info bubble: There are a number of diagrams and figures which illustrates TCP/IP's layered design and show the relationships among the core protocols. As a unit of data flows downward from a network application to the network interface card, that data travels through a succession of TCP/IP modules. At each step along the way, it is packaged with information required by the equivalent TCP/IP module on the other end. By the time the data makes it to the network card, it's a standard *Ethernet* frame, assuming the network is an Ethernet network. The TCP/IP software on the receiving end recreates the original data for the receiving application by grabbing the Ethernet frame and passing it upward through the TCP/IP stack.

To picture what role TCP/IP plays in real-world networks, consider what happens when a *Web browser* uses *HTTP (HyperText Transfer Protocol*) to retrieve a page of *HTML* data from a *Web server* attached to the Internet. The browser uses a high-level software abstraction called a socket to form a virtual connection to a server. To retrieve a Web page, it sends an HTTP GET command to the server by writing the command to the socket. The socket software in turn uses TCP to send the bits and bytes comprising the GET command to the Web server. TCP segments the data and passes the individual segments to the IP module, which transmits the segments in datagrams to the Web server. If the browser and the server are running on computers connected to different physical networks (as is usually the case), the datagrams go from network to network until they reach the one to which the server is physically connected. Ultimately, the datagrams are delivered to their destination and reassembled so that the Web server, which reads chunks of data by performing reads on its socket, gets a continuous stream of data. To the browser and the server, data written to the socket at one end shows up at the other end, as if by magic. But underneath, all sorts of complex interactions have taken place to create an illusion of seamless data transfer across networks. That is what TCP/IP is all about: turning lots of small networks into one big network and providing the services that applications need to communicate with each other over the resulting Internet.

Ethernet Info Bubble:
Ethernet is a local area network (LAN) developed by Xerox, Digital Equipment and Intel. It is governed by the IEE 802.3 standard and is the most widely used LAN architecture today. Different flavors of Ethernet include think Ethernet, ThinNet, or CheaperNet, and Twisted Pair Ethernet (10BaseT). Also Fast Ethernet is starting to become very popular. Think and Thin Ethernet are rarely installed anyomre and are more often then not being upgraded to Twisted Pair Ethernet or Fast Ethernet (100BaseT). Twisted Pair Ethernet executes across the cable at 10Mbit/second and Fast Ethernet runs across the cable at 100Mbit/second.

NT-NetWare Coexistence info bubble:

- NWLink runs on an NT server and provides an IPX/SPX-compatible transport layer. It allows NetWare clients to access NT server-based applications over IPX without changing client software. Without an NT-based IPX stack, NetWare clients would have to run dual protocol stacks and access NT-based applications over TCP/IP.
- Client Service For NetWare allows an NT workstation to access NetWare file and print services and many NetWare server-based applications over IPX or TCP/IP.
- Gateway Service For NetWare allows an NT server to act as a gateway to NetWare for dial-in or TCP/IP-based clients.
- File and Print Service For NetWare makes an NT server look like a NetWare 3.12 file and print server, so NetWare clients can use NT as a file, print, and application server without changing client software.

- Directory Service Manager For NetWare copies NetWare user and group account information from NetWare 2.x and 3.x servers to NT servers and then incrementally propagates any changes made to the accounts in the NT environment back to the NetWare servers.
- Migration Tool For NetWare automatically migrates NetWare 2.x, 3.x, and 4.x user and group accounts, log-in scripts, and security setups to an NT server.

More Resources Information Bubble:

If you're interested in learning more about TCP/IP, we recommend two further sources of information. The first is RFC (Request for Comment) 1180, a document entitled; *"A TCP/IP Tutorial"*. This is a public document which you can download from any Internet site that disseminates RFCs. The second source is the book *"Internetworking with TCP/IP, Volume 1: Principles, Protocols, and Architectures"*, by Douglas E. Comer (1995, Prentice-Hall). The first in a three-part series, it's considered by many to be the TCP/IP bible.

Publishers are swamping the trade magazines with articles and bookstores with books about the Internet. This list captures some that we have found personally useful and which will help you get oriented to Web servers, HTML, and electronic publishing. This list is not meant to be exhaustive or to reflect unfavorably on other books or articles covering the same material.

Non-Technical Overviews:

- Connecting to the Internet, by Susan Estrada. Sebastopol, CA: O'Reilly & Associates, 1993. ISBN: 1-56592-061-9.
- "Mosaic--Beyond Net Surfing," by John R. Vacca. BYTE, January 1995,75ff.
- "The Second Phase of the Revolution Has Begun," by Gary Wolf. Wired, October 1994, pp. 116ff.
- "The Web Untangled," by Rick Ayre and Kevin Richard. PC Magazine, February 7, 1995, pp. 173ff.
- "Web Watch," by Jack Rickard. Boardwatch, December 1994, pp. 40ff.
- The Whole Internet User's Guide & Catalog, Second Edition, by Ed Krol.
- Chapter 13: "The World-Wide-Web." Sebastopol, CA: O'Reilly and Associates, 1994. ISBN: 1-56592-063-5.
- Zen and the Art of the Internet: A Beginner's Guide to the Internet, Third Edition, by Brendan Kehoe. Englewood Cliffs, NJ: Prentice Hall, ISBN: 0-13-121492-6.

Technical Overviews:

- "Getting Wired into the Internet: A Crash Course on FTP, Gopher, Web,and More," by J. Allard and Steven Sinofsky. Microsoft Systems Journal, September 1994, pp. 53ff.
- "Mosaic and the World-Wide Web," by Robert J. Vetter, Chris Spell, and Charles Ward. Computer, October 1994, pp. 49ff.
- "NCSA Mosaic and the World-Wide-Web: Global Hypermedia Protocols for the Internet," by Bruce R. Schatz and Joseph B. Hardin. Science, August 12, 1994, pp. 895ff.
- "The World-Wide Web," by Tim Berners-Lee, Robert Cailliau, Ari Luotonene, Henrik Frystyk Nielsen, and Arthur Secret. Communications of the ACM, August 1994, pp. 76ff.

Web Server Management and HTML Authoring:

- HTML Manual of Style, by Larry Aronson. Emeryville, CA: Ziff-Davis Press, 1994. ISBN: 1-56276-300-8.
- Managing Internet Information Resources, by Cricket Liu, Jerry Peek, Russ Jones, Bryan Buus, and Adrian Nye. Sebastopol, CA: O'Reilly and Associates, 1994. ISBN: 1-56592-062-7.
- "World-Wide-Web and HTML," by Douglas C. McArthur. Dr. Dobb's Journal, December 1994, pp. 18ff.

Once you have an Internet connection and a working ***Web browser***, there is a huge number of overviews, tutorials, specifications, and other helpful resources available on the Web itself. You can locate these

resources most efficiently by going to the World-Wide Web home page at CERN (http://info.cern.ch), selecting the pages related to *Web servers*, browsers, and authoring tools, and following the *hyperlinks* from there.